Early-onset Wilson disease caused by *ATP7B* exon skipping associated with intronic variant

Daniel C. Koboldt,[1,2] Scott E. Hickey,[1,3] Bimal P. Chaudhari,[1,2,3,4]
Theresa Mihalic Mosher,[1,2,3] Tracy Bedrosian,[2] Erin Crist,[1,2]
Stephen G. Kaler,[1,3,5] Kim McBride,[1,3,6] Peter White,[1,3]
and Richard K. Wilson[1,3]

[1]Department of Pediatrics, The Ohio State University, Columbus, Ohio 43205, USA; [2]Institute for Genomic Medicine at Nationwide Children's Hospital, Columbus, Ohio 43205, USA; [3]Division of Genetic and Genomic Medicine, [4]Division of Neonatology, Nationwide Children's Hospital, Columbus, Ohio 43205, USA; [5]Center for Gene Therapy, [6]Center for Cardiovascular and Pulmonary Research, Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, Ohio 43205, USA

Abstract Wilson disease is a medically actionable rare autosomal recessive disorder of defective copper excretion caused by mutations in *ATP7B*, one of two highly evolutionarily conserved copper-transporting ATPases. Hundreds of disease-causing variants in *ATP7B* have been reported to public databases; more than half of these are missense changes, and a significant proportion are presumed unequivocal loss-of-function variants (nonsense, frameshift, and canonical splice site). Current molecular genetic testing includes sequencing all coding exons (±10 bp) as well as deletion/duplication testing, with reported sensitivity of >98%. We report a proband from a consanguineous family with a biochemical phenotype consistent with early-onset Wilson disease who tested negative on conventional molecular genetic testing. Using a combination of whole-genome sequencing and transcriptome sequencing, we found that the proband's disease is due to skipping of exons 6–7 of the *ATP7B* gene associated with a novel intronic variant (NM_000053.4:c.1947-19T > A) that alters a putative splicing enhancer element. This variant was also homozygous in the proband's younger sister, whose subsequent clinical evaluations revealed biochemical evidence of Wilson disease. Our work adds to emerging evidence that *ATP7B* exon skipping from deep intronic variants outside typical splice junctions is an important mechanism of Wilson disease; the variants responsible may elude standard genetic testing.

Corresponding author:
Daniel.Koboldt@
nationwidechildrens.org

[Supplemental material is available for this article.]

INTRODUCTION

High-throughput sequencing (HTS) technologies have dramatically accelerated the pace of human disease genetics (Koboldt et al. 2013). Exome and genome sequencing have elucidated the genetic basis for thousands of rare inherited disorders in the past decade. HTS has also supplanted other genomic interrogation technologies for the vast majority of clinical genetic tests; for patients with rare disorders of likely genetic etiology, clinical exome sequencing is often a frontline diagnostic test. Despite this progress, a recent meta-analysis of 37 studies comprising 20,068 children found that the diagnostic utility of exome sequencing (ES) and whole-genome sequencing (WGS) is only 36%–41% (Clark et al. 2018). That

proportion has changed little despite the continued identification of novel disease genes. It is likely that an appreciable number of undiagnosed cases harbor pathogenic variants in known disease genes that are refractory to identification by current clinical genetic testing paradigms. Patients who are clinically diagnosed with a disease of well-established etiology but whose clinical genetic test results are negative offer an opportunity to characterize these "occult" pathogenic variants using new genomic and analytic techniques.

Wilson disease (WD) is a rare recessive disorder characterized by a dramatic accumulation of intracellular copper—primarily in the liver and brain—as a result of defective copper excretion (de Bie et al. 2007). It was first described more than a century ago by neurologist S.A. Kinnier Wilson, who called it *progressive lenticular degeneration*, a "familial nervous system disease" involving cirrhosis of the liver (Wilson 1912; Bandmann et al. 2015). Clinical manifestations include not only hepatic abnormalities (e.g., cirrhosis) but also neurologic and psychiatric symptoms. Clinical diagnosis is often possible before the appearance of neurologic symptoms through biochemical testing for reduced serum ceruloplasmin (assay-dependent cutoff generally <20 mg/dL) and elevated 24-h urinary copper (>40 mcg) (Roberts et al. 2008). The worldwide incidence of WD is estimated at 1:30,000 to 1:100,000, although incidence as high as 1:7000 has been reported in founder populations (Loudianos et al. 1999).

WD is caused by biallelic variants in the *ATP7B* gene at Chr 13q14.3 (Bull et al. 1993; Petrukhin et al. 1993; Tanzi et al. 1993). This gene encodes a metal-transporting P-type adenosine triphosphatase (ATPase) whose function is in the transmembrane transport of copper within hepatocytes, ultimately shuttling copper for excretion into the bile. Absence, reduction, or decreased function of the ATP7B protein leads to hepatic copper accumulation and injury (Roberts et al. 2008; Bandmann et al. 2015). *ATP7B* is highly homologous with *ATP7A*, an X-linked gene in which mutations cause Menkes disease, an essentially inverse phenotype of generalized copper deficiency (Kaler 1998). Both *ATP7B* and *ATP7A* encode evolutionarily conserved transmembrane proteins with metal-binding and ATPase domains (de Bie et al. 2007; Petrukhin et al. 1994). Some 192 variants in *ATP7B* have been reported as pathogenic or likely pathogenic to the ClinVar database; of these, loss-of-function variants including frameshift ($n = 72$), nonsense ($n = 35$), and splice site ($n = 29$) variants together represent >70% of disease-causing alterations. More than 500 disease-causing variants in *ATP7B* were reported to the Wilson Disease Mutation Database (2001–2010) (Kenney and Cox 2007). Although more than half such variants are missense changes, appreciable proportions of loss-of-function variants (nonsense, frameshift, and splice site) were reported.

The clinical testing method to identify *ATP7B* variants uses a combination of next-generation sequencing (NGS) and Sanger sequencing technologies to cover the full coding regions of *ATP7B* plus ~10 bases on noncoding DNA flanking each exon. Copy-number variants of four exons or more in size are detected through analysis of NGS data with sensitivity approaching 100%. The clinical sensitivity of this approach is listed as ~98% among individuals with WD (Prevention Genetics; https://www.preventiongenetics.com/testInfo?val=Wilson+Disease+%2F+Hepatolenticular+Degeneration+via+the+ATP7B+Gene).

We report a proband from a consanguineous family who carried a clinical and biochemical diagnosis of WD but whose clinical molecular genetic testing was negative. Through a combination of WGS, RNA sequencing, and computational analysis, we determined that the disease is caused by exon skipping in *ATP7B* associated with a novel intronic variant. Our analysis also revealed that the proband's 7-yr-old younger sister, who was clinically asymptomatic, was homozygous for this variant. Her subsequent clinical workup revealed biochemical evidence of WD consistent with the molecular genetic findings. Our work adds to emerging evidence that exon skipping in *ATP7B* represents a common etiology for WD and illustrates the value of applying new genomic interrogation techniques to identify occult pathogenic variants.
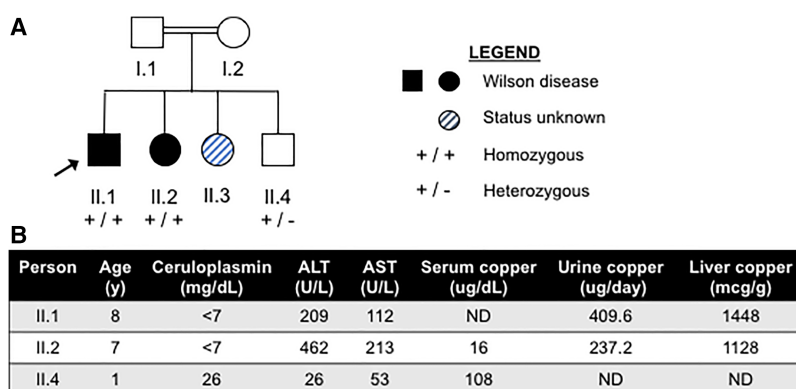
Figure 1. Clinical and family history. (*A*) Family pedigree with disease status (shading) and genotypes for the c.1947-19A > T variant, if known. The proband is indicated with an arrow. (*B*) Biochemical test results for three children who were tested. (ALT) Alanine transaminase, (AST) aspartate transaminase, (ND) not done.

RESULTS

Clinical Presentation and Family History

The proband is an 8-yr-old Palestinian male who presented to the Nationwide Children's Hospital Gastroenterology Clinic shortly after he and his father came to the United States. The family was evaluated with an Arabic interpreter (pedigree in Fig. 1A). The proband had been diagnosed with WD by a physician in Jordan at 5 yr of age. He reportedly was receiving oral penicillamine at the time of his initial examination in the United States. The family noted that the diagnosis of WD had been established based on pallor, hepatomegaly, and findings on eye exam (presumed to be Kayser–Fleischer rings). We were unable to contact the clinician who made the original clinical diagnosis and the patient presented without pertinent clinical records.

Evaluations for the proband at our hospital included an abdominal ultrasound that showed severe hepatomegaly with steatosis and mild splenomegaly. On laboratory evaluation (Fig. 1B), serum transaminases were mildly elevated (ALT 209/AST 112) and serum ceruloplasmin was undetectable. The 24-h urine copper was markedly elevated (409.6). Repeat eye exam noted right conjunctival nevus but detected no Kayser–Fleischer rings. However, the proband was diagnosed at a relatively young age and started chelation therapy in his home country, which may explain the absence of corneal copper deposits. A liver needle biopsy showed marked panlobular steatosis, with moderate necroinflammatory activity present, ballooning cell injury and cirrhosis, and elevated copper (1448 mcg/g; normal ≤35 mcg/g).

Commercial *ATP7B* sequencing with copy-number variant (CNV) detection performed by Prevention Genetics (test code 7871) reported no sequence variants likely to be a primary cause of disease. Single-nucleotide polymorphism (SNP) chromosomal microarray was consistent with ∼8.96% homozygosity across the autosomal genome. A run of homozygosity (ROH) within the 13q14.3q31.3 region encompassing *ATP7B* was noted. A family pedigree showed three younger full siblings whose disease status was unknown. Two siblings reside in the United States; the third resides overseas and was unavailable for testing.

Genomic Analyses

Given the clinical diagnosis of Wilson disease and ROH at Chr 13q14.3 noted from array testing, we suspected that an occult variant in *ATP7B* was the likely cause of disease. We therefore elected to perform WGS on peripheral blood from proband (II.1) and available family

members (both parents, I.1 and I.2, and his younger sister, II.2; Fig. 1A). At the time, the sister (II.2) was presumed healthy and had not undergone extensive clinical evaluations. We also performed RNA-seq on frozen liver tissue from the proband, which had been obtained as part of the clinical workup, and requested the raw sequence data for the *ATP7B* clinical testing from Prevention Genetics.

Genome Sequencing

We performed WGS on the four family members on an Illumina NovaSeq6000 instrument (2 × 150 bp), generating ~162 Gbp per sample on average. Mapping the sequence reads to the GRCh37 reference sequence yielded ~45.4× haploid coverage per sample (range 43.1× to 47.5×; see Table 1 for metrics).

Genome and Variant Analysis

Consistent with previous testing, our analysis of WGS data for the proband revealed ROH encompassing ~9.73% of the genome, including a ~3.93-Mbp segment at Chr 13q14.3-q31.3 (Fig. 2A) that covers the *ATP7B* locus. Multisample variant calling in the family quad uncovered 155 small variants (SNVs/indels) that were annotated in or near *ATP7B* by the Variant Effect Predictor (VEP 90). Of these, the vast majority ($n = 127$) had a minor allele frequency of >0.05 in the gnomAD database (Lek et al. 2016), making them too common to cause a recessive disorder. Another five variants were manually reviewed by trained reviewers using the Integrative Genomics Viewer and were deemed to represent local alignment artifacts (misalignment of reads to the reference sequence) rather than true variation. The 23 variants passing these filters are detailed in Supplemental Table 1. Of these, 20 had global minor allele frequency (MAF) > 0.01 and at least two homozygotes in the gnomAD database and are thus unlikely to cause disease. One variant was a small homopolymer insertion in a low-complexity repeat region, and one variant with somewhat low MAF

Table 1. Sequencing metrics

	Proband	Mother	Father	Sibling
RNA sequencing (liver)				
Total reads (millions)	250.6			
Unmapped	3.31%			
HQ mapped	88.27%			
LQ mapped	7.31%			
Ribosomal mapped	1.12%			
Coding/UTR reads (millions)	138.4			
Coding/UTR fraction	62.56%			
Whole-genome sequencing (blood)				
Total sequence (Gbp)	168	159	156	164
Mean read length (bp)	145.2	145.2	145.2	145.1
Map rate (b37)	97.86%	97.97%	96.96%	97.87%
Duplication rate	6.61%	7.04%	6.47%	6.59%
Median insert size (bp)	340	346	326	341
Average depth	47.5×	45.1×	43.1×	45.7×

(Top) RNA sequencing metrics for proband liver tissue obtained via biopsy for clinical purposes. (Bottom) Whole-genome sequencing metrics for genomic DNA from peripheral blood. All samples were sequenced on Illumina NovaSeq instruments (2 × 150 bp).
(HQ) High quality (mapping quality > 0), (LQ) low quality (mapping quality = 0), (UTR) untranslated region.
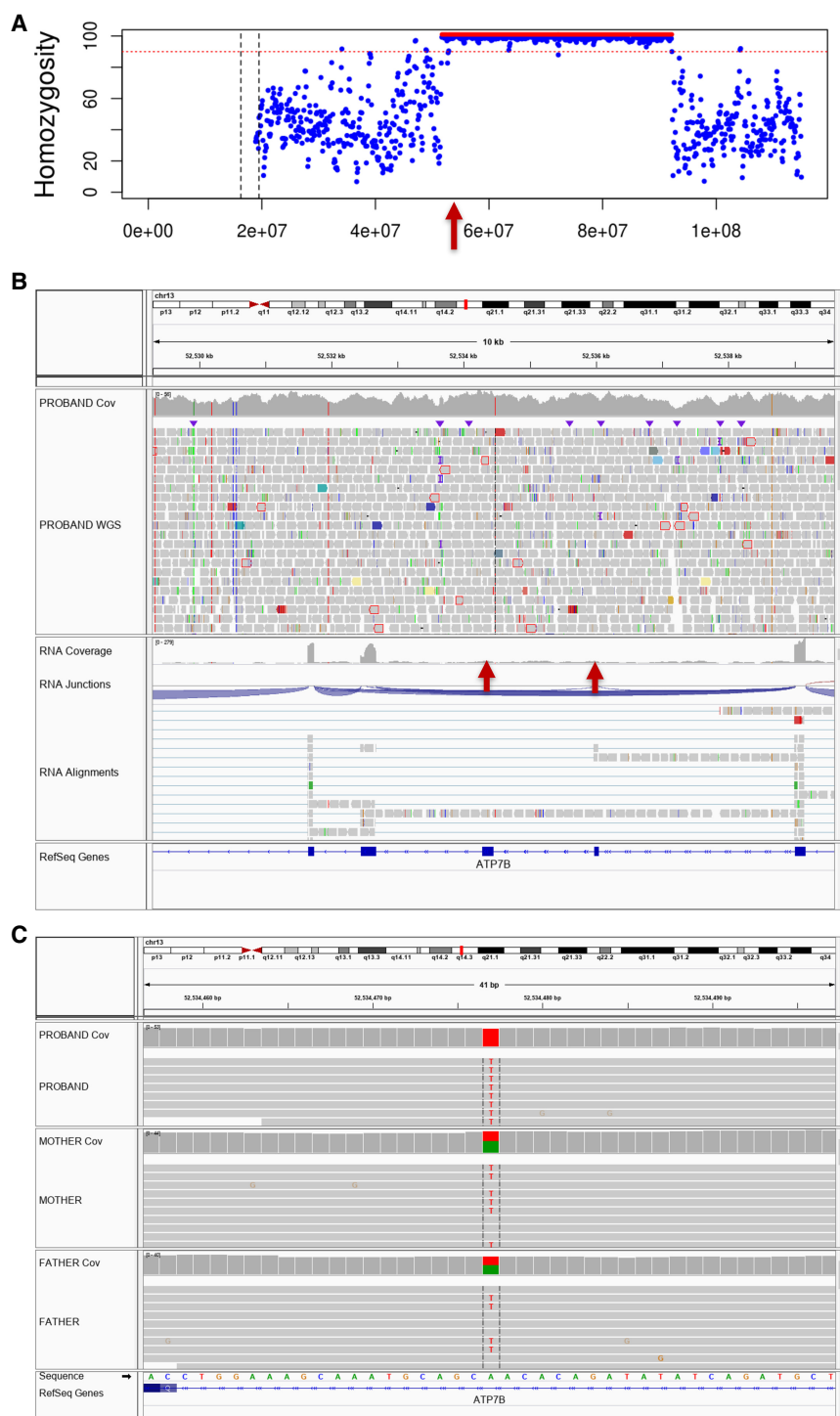
Figure 2. Genomic analysis of case RD190073. (*A*) Run of homozygosity (ROH) analysis using VarScan v2.3.4 reveals a large ROH region on Chromosome 13 overlapping the *ATP7B* gene, indicated by a red arrow. (*B*) Integrative Genomics Viewer (IGV) screenshot of proband genome sequence (*top* track) and transcriptome sequence (*bottom* track). Exons 6 and 7 are not expressed in the RNA (red arrows). (*C*) IGV screenshot of the c.1947-19T > A variant showing its proximity to the exon and zygosity (homozygous in the proband, heterozygous in both parents).

(rs558801757, MAF = 0.0041) is located more than 1700 bp downstream from the ATP7B region and thus deemed unlikely to affect splicing.

The remaining variant was an A-to-T change in the sixth intron of *ATP7B* (NM_000053.4: c.1947-19T > A). It is not present in the gnomAD database (120,839 individuals had coverage). This suggests that the variant is extremely rare, although it should be noted that Middle Eastern populations are underrepresented in public sequence databases. Of note, this variant is close enough to the nearby exon that it achieved good coverage and was clearly visible in the panel sequencing data (184× depth). However, it was not reported by the testing provider, likely because it lies well beyond the canonical splice site and does not appear to be conserved (Genomic Evolutionary Rate Profiling-rejected substitutions [GERP-RS] score = −0.68, DANN score = 0.8337).

RNA Sequencing Reveals Exon 6–7 Skipping

We next examined expression data from the proband's liver tissue. As expected, *ATP7B* is highly expressed in liver tissue; we achieved ~200×–300× read depth across most of the gene. However, read depth was <35× in exons 6 and 7 despite 250×–300× depth at the adjacent exons (5 and 8), suggesting absence of these exons in processed transcripts (Fig. 2B, bottom track). In support of this, at least 22 exon-spanning reads (blue arcs) map from exon 5 to exon 8, with no sequence aligning to exons 6 and 7. There were 13 RNA-seq reads present at the intronic position c.1947-19T > A, likely reflecting unprocessed mRNA transcripts. All 13 contained the variant allele, which offers independent molecular confirmation that the variant is real and not a sequencing or alignment artifact.

Because the c.1947-19T > A variant was the only plausible genomic source of the splice disruption, we analyzed the variant and its sequence context in detail with two independent tools, Human Splicing Finder (HSF) v3.1 (Desmet et al. 2009) and RBPmap v1.1 (Paz et al. 2014). Using the reference or variant allele with ±30 bp of flanking sequence in each direction, we identified two factors that showed strikingly different results between reference and variant alleles (Table 2). First, we found that the variant allele disrupts a predicted binding motif of SC35 (SRSF2), a splice enhancer, reducing its HSF score from 77.01 to 61.89 (below the default threshold of 75.05). Similarly, RBPmap predicts a high-scoring SC35 motif for the reference allele (Z = 2.495, P = 0.0063) but a low-stringency match (Z = 1.949, P = 0.0256) for the variant allele. Second, we found that the variant allele creates a high-scoring motif for heterogeneous nuclear ribonucleoprotein (hnRNP) A1, a splicing silencer (HSF score = 70.24, threshold = 65.476). RBPmap supports these results, predicting a high-scoring hnRNP A1 motif for the variant allele (Z = 2.623, P = 0.00436) but no matching motif for the reference allele. These computational results predict that the c.1947-19T > A variant disrupts a splice enhancer while creating a splicing silencer motif, results consistent with the aberrant splicing pattern we observed. The variant's position between the two affected exons

Table 2. Results of splice motif analysis for variant position ±30 bp

Position	Motif	Factor	HSF Ref Score	HSF Alt Score	RBP Ref Z-score	RBP Alt Z-score
−22	TGT[T/A]GCTG	SC35 (SRSF2)	77.01	61.89*	2.495	1.949*
−20	T[T/A]GCTG	hnRNP A1	34.52*	70.24	n/a*	2.623

Shown are the scores for two factors whose motifs are affected using either the reference (Ref) or alternate (Alt) allele from Human Splicing Finder (HSF) and RBPmap (RBP). Scores marked with (*) are below the default threshold for the factor.
(hnRNP) Heterogeneous nuclear ribonucleoprotein.

and the lack of any other viable candidate variants in the region offer compelling evidence that the c.1947-19T > A variant is the cause of the aberrant splicing.

Impact of Exon 6–7 Loss

Exon 6 (77 bp) and exon 7 (175 bp) are normally expressed in ATP7B in liver tissue, and together encode precisely 84 amino acids (residues 624–707) of the protein (Fig. 3A). By analogy to *ATP7A*, exon 6 of *ATP7B* encodes a short cytosolic segment linking the molecule's copper-binding region (six recurring CXXC motifs) to the initial transmembrane (TM) segment (Kaler et al. 1995; Kaler 2011). Exon 7 is predicted to encode the first two transmembrane helices attached by a luminal loop. Although deletion of these exons would not disrupt the proper *ATP7B* translational reading frame, it would severely impair the mutant protein's copper transport capacity. The first two transmembrane segments normally enable formation of a copper "platform" as determined by X-ray crystallography of the homologous *Legionella pneumophila* CopA copper ATPase (Gourdon et al. 2011). This crystal structure indicates a three-stage copper transport pathway involving several conserved residues including a transmembrane kink at a double-glycine motif displaying an amphipathic helix lining a putative copper entry point. The double glycine link is specific for P-type ATPases of the class IB including ATP7A (Yi and Kaler 2015) and occurs in exon 8 of *ATP7B*, within two residues of the exon 7 boundary. In the absence of exons 6 and 7, this double-glycine motif would mislocalize to the cytosolic aspect of intracellular compartments, impeding proper transmembrane copper passage. Analogous exon skipping has been reported in classic Menkes disease patients with an *ATP7A* splice junction variant that generated multiple mutant transcripts, none of which appeared functional (Kaler et al. 1995).
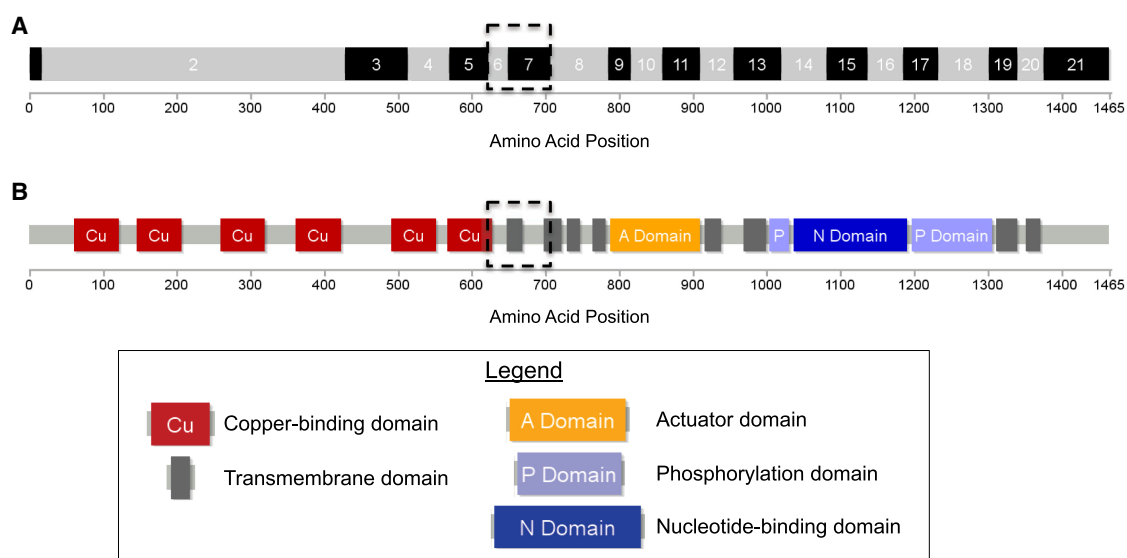

Figure 3. Impact of aberrant splicing on ATP7B structure. (*A*) Exon structure of ATP7B, with even and odd numbered exons differently colored for distinction. (*B*) Protein structure of ATP7B showing key structural elements and functional domains. Region boundaries were obtained from PFAM (Cu-binding domains) or the published crystal structure of ATP7B (all others). Skipping of exons 6 and 7 (dotted rectangle) removes 84 amino acids (residues 624–707), removes the first transmembrane domain (often called TMA), and may also impact the sixth Cu-binding domain (MBD6) and second transmembrane domain (TMB).

Table 3. Genomic findings and variant interpretation

Gene	Variant	HGVS	Proband	Mother	Father	Sibling	Interpretation
ATP7B	Chr 13: 52534477 A/T (hg19)	NM_000053.4: c.1947-19T > A	Hom	Het	Het	Hom	LP (PS3, PM2, PP1, PP4)

Genomic coordinates reflect build GRCh37.
(HGVS) Human Genome Variation Society, (Hom) homogeneous, (Het) heterogeneous, (LP) likely pathogenic.

Evaluation of Family Members

Although the proband's 7-yr-old sister was described as healthy at the time of enrollment, genome sequencing revealed that she, too, had a large ROH encompassing the *ATP7B* gene region (Supplemental Fig. 1) and was homozygous for the c.1947-19T > A variant (Table 3). Based on this finding, she underwent clinical evaluation at our hospital. She had elevated transaminases (ALT 173 U/l, AST 373 U/l), low serum ceruloplasmin, and high urinary copper excretion (Fig. 1B). She had normal albumin (4.8 g/dL), alkaline phosphatase (278 U/L), and bilirubin (0.6 mg/dL). Complete blood count (CBC) showed normal values for white blood cells (WBCs) $6.0 \times 10^3/\mu L$, hemoglobin (Hgb) 13.7 g/dL, hematocrit (Hct) 39.2%, and platelet count (Plt) 362 × 103/μL. International normalized ratio (INR) and prothrombin time (PT) were not elevated at 0.97 and 12.9 sec, respectively. Her liver copper was elevated (1128 mcg/g; normal ≤ 35 mcg/g), confirming a diagnosis of Wilson disease. Thus, our genomic analysis enabled early diagnosis of a second family member.

Variant Interpretation

Variant interpretation under American College of Medical Genetics and Genomics (ACMG) recommended guidelines is challenging for this variant. It is absent from public databases (PM2) and segregates with disease in the proband under the expected recessive inheritance pattern (PP1), and the phenotype is specific for a disease with a single known genetic etiology (PP4). This genomic evidence alone is insufficient to make a classification other than variant of uncertain significance (VUS). However, we would argue that the aberrant splicing observed in the proband's liver tissue (i.e., skipping of exons 6–7), together with our computational analysis, represents in vivo functional evidence that the variant disrupts the ATP7B gene product. In this view, the PS3 criterion would also apply, which would shift the variant's classification to likely pathogenic (Table 3).

DISCUSSION

WD is a medically actionable autosomal recessive disorder of copper metabolism with variable age of clinical presentation. *ATP7B* is the only gene known to cause WD. The phenotype of WD is clinically distinct from other disorders of copper metabolism such as Menkes disease and MEDNIK syndrome (Umair and Alfadhel 2019). Considering this, as well as the fact that the SNP ROH encompassed the *ATP7B* gene, a hidden variant involving *ATP7B* was suspected. WGS identified a homozygous deep intronic variant that is predicted to cause exon skipping by splicing computational tool. RNA-seq data showed 90% lower expression of exons 6 and 7, which is consistent with exon skipping. Removing amino acids 624–707 of the ATP7B protein would affect metal binding domain (MBD6) which would be expected to disrupt copper transport activity (Shanmugavel et al. 2019). The c.1947-19T > A is the only plausible genomic variant in this region to underlie the aberrant splicing. Because homozygosity for this variant segregates with WD in the family, further support of disease causality is

evident. Analysis of the Δ exon 6–7 *ATP7B* transcript by yeast complementation (Forbes and Cox 1998; Kaler et al. 2008) could provide a formal laboratory-based approach to confirm loss of copper transport function, although this was not completed given the multiple other lines of supporting evidence.

Several hundred disease-causing variants in *ATP7B* have been reported to the Wilson Disease Mutation Database (Kenney and Cox 2007), the Human Gene Mutation Database, and the ClinVar database. Less than half are classified as unequivocal loss-of-function variants (nonsense, frameshift, or splice site). Canonical splice site variants presumably disrupt splicing and are generally considered pathogenic, but the impact of variants at other positions in the splice region is less certain (Dong et al. 2016). However, there is emerging evidence that variants outside of the canonical splice site can disrupt ATP7B function via effects on splicing. Wang et al. used a minigene assay to identify both missense and synonymous coding variants that caused exon skipping in *ATP7B* (Wang et al. 2018). Liu et al. used a similar approach to show that a variant in intron 4 (c.1707 + 5G > A) causes exon 4 skipping (Liu et al. 2018). To our knowledge, the c.1947-19T > A variant in our family is the deepest intronic variant to be implicated in *ATP7B* exon skipping to date.

Clinical recognition of WD can be difficult as there is no single specific clinical or biochemical finding pathognomonic for the diagnosis. In equivocal situations, liver biopsy for copper quantification can be useful but has the drawback of being invasive and may be negative in early stages of WD. Molecular genomic testing is an extremely useful addition to the diagnostic armamentarium; however, as in the case described here, some disease-causing variants may not be detected by standard clinical genetic testing.

METHODS

Clinical Workup

Serum ceruloplasmin levels for the proband and his sibling were measured by nephelometric immunoassay on a Siemens BNII nephelometer. Urine copper was measured by inductively coupled plasma mass spectrometry (ICP-MS). Ophthalmologic evaluation of the proband included a fundus examination and slit-lamp exam. Commercial genetic testing was performed by Prevention Genetics (test code 7871, "ATP7B gene sequencing with CNV detection") using genomic DNA from proband whole blood. According to the report, the test achieved 325× average NGS coverage, with 100% of bases covered (≥20× for exons and ±10 bp of flanking DNA). The report of negative findings was received in February 2019.

Library Preparation

Following consent and enrollment, genomic DNA from the proband, parents, and younger sister was extracted from peripheral blood. Paired-end genome sequencing libraries were constructed using NEBNext Ultra II FS DNA Library Prep Kit (New England BioLabs) according to manufacturer protocols. RNA was extracted from a previously frozen liver needle biopsy specimen obtained as part of the proband's clinical workup; sequencing libraries were constructed using NEBNext Ultra II RNA Library Prep Kit (New England BioLabs) according to manufacturer protocols.

RNA Sequencing and Analysis

RNA from the liver specimen underwent DNase treatment and ribodepletion prior to library construction using Illumina's TruSeq Stranded total RNA kit. RNA sequencing was performed using a NovaSeq6000 instrument (2 × 150 bp) according to manufacturer protocols generating 150.6 million reads. The default RNA-seq pipeline aligned reads to the GRCh37

(hg19) assembly of the human using version 2.6.0c of the RNA-Seq aligner STAR (Dobin et al. 2013). Transcript features were identified from the GFF file provided with the assembly from NCBI and raw coverage counts were calculated using HTSeq (Anders et al. 2015) v0.10.0.

Genome Sequencing and Analysis

WGS was performed on genomic DNA from the proband, mother, father, and two full siblings (II.2 and II.4) using an Illumina NovaSeq6000 instrument according to manufacturer protocols. Reads were mapped to the GRCh37 reference sequence and secondary data analysis was performed using Churchill (Kelly et al. 2015), which implements the GATK "best practices" workflow for alignment, variant discovery and genotyping. The average sequence depth achieved was ~45.4× for the blood samples (proband, mother, father, and sibling); detailed sequencing metrics are provided in Table 1. Variants were joint-called in all family members using GATK 4.0.5.1, and the resulting VCF file was annotated with gene, transcript, function class, damaging scores, and population allele frequencies using an in-house pipeline built around the SNPeff (Cingolani et al. 2012) annotation tool. ROH analysis was performed using VarScan (Koboldt et al. 2012) v2.4.3.

Our general approach to variant annotation and prioritization has already been described (Koboldt et al. 2018). After removing common variants (MAF > 0.01 in the gnomAD database), we selected for further analysis all splice site, frameshift, and nonsense variants, as well as missense variants predicted to be damaging by SIFT (score < 0.05), PolyPhen (score > 0.453), GERP (score > 2.0), or CADD (Phred score > 15). Because of the family history, we prioritized rare variants segregating with disease under a recessive inheritance model (i.e., homozygous in the proband, heterozygous in both parents). Other inheritance models—dominant de novo, compound heterozygous, and X-linked—were considered but did not produce compelling candidate variants in the gene associated with WD.

ADDITIONAL INFORMATION

Data Deposition and Access

The variant and our interpretation have been submitted to ClinVar (https://www.ncbi.nlm.nih .gov/clinvar/) under accession number VCV000812709.1. Deposition of raw sequencing data was not permitted because of lack of patient consent.

Ethics Statement

Written informed consent was obtained for all participants in this study under a research protocol approved by the Institutional Review Board at Nationwide Children's Hospital (IRB11-00215 Study: Using Genome Sequencing to Identify Causes of Rare Birth Defects and Rare Disorders).

REFERENCES

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31:** 166–169. doi:10.1093/bioinformatics/btu638

Bandmann O, Weiss KH, Kaler SG. 2015. Wilson's disease and other neurological copper disorders. *Lancet Neurol* **14:** 103–113. doi:10.1016/S1474-4422(14)70190-5

Bull PC, Thomas GR, Rommens JM, Forbes JR, Cox DW. 1993. The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat Genet* **5:** 327–337. doi:10.1038/ng1293-327

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: sNPs in the genome of *Drosophila melanogaster* strain w^{1118}; *iso-2*; *iso-3. Fly (Austin)* **6:** 80–92. doi:10.4161/fly.19695

Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, Kingsmore SF. 2018. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med* **3:** 16. doi:10.1038/s41525-018-0053-8

Compston A. 1912. Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver, by S.A. Kinnier Wilson (From the National Hospital, and the Laboratory of the National Hospital, Queen Square, London) (Brain 1912: 34; 295–509) *Brain* **132:** 1997–2001. doi:10.1093/brain/awp193

de Bie P, Muller P, Wijmenga C, Klomp LW. 2007. Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes. *J Med Genet* **44:** 673–688. doi:10.1136/jmg.2007.052746

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* **37:** e67. doi:10.1093/nar/gkp215

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29:** 15–21. doi:10.1093/bioinformatics/bts635

Dong Y, Ni W, Chen WJ, Wan B, Zhao GX, Shi ZQ, Zhang Y, Wang N, Yu L, Xu JF, et al. 2016. Spectrum and classification of ATP7B variants in a large cohort of Chinese patients with Wilson's disease guides genetic diagnosis. *Theranostics* **6:** 638–649. doi:10.7150/thno.14596

Forbes JR, Cox DW. 1998. Functional characterization of missense mutations in *ATP7B*: Wilson disease mutation or normal variant? *Am J Hum Genet* **63:** 1663–1674. doi:10.1086/302163

Gourdon P, Liu XY, Skjorringe T, Morth JP, Moller LB, Pedersen BP, Nissen P. 2011. Crystal structure of a copper-transporting PIB-type ATPase. *Nature* **475:** 59–64. doi:10.1038/nature10191

Kaler SG. 1998. Metabolic and molecular bases of Menkes disease and occipital horn syndrome. *Pediatr Dev Pathol* **1:** 85–98. doi:10.1007/s100249900011

Kaler SG. 2011. ATP7A-related copper transport diseases—emerging concepts and future trends. *Nat Rev Neurol* **7:** 15–29. doi:10.1038/nrneurol.2010.180

Kaler SG, Buist NR, Holmes CS, Goldstein DS, Miller RC, Gahl WA. 1995. Early copper therapy in classic Menkes disease patients with a novel splicing mutation. *Ann Neurol* **38:** 921–928. doi:10.1002/ana.410380613

Kaler SG, Holmes CS, Goldstein DS, Tang J, Godwin SC, Donsante A, Liew CJ, Sato S, Patronas N. 2008. Neonatal diagnosis and treatment of Menkes disease. *N Engl J Med* **358:** 605–614. doi:10.1056/NEJMoa070613

Kelly BJ, Fitch JR, Hu Y, Corsmeier DJ, Zhong H, Wetzel AN, Nordquist RD, Newsom DL, White P. 2015. Churchill: an ultra-fast, deterministic, highly scalable and balanced parallelization strategy for the discovery of human genetic variation in clinical and population-scale genomics. *Genome Biol* **16:** 6. doi:10.1186/s13059-014-0577-x

Kenney SM, Cox DW. 2007. Sequence variation database for the Wilson disease copper transporter, ATP7B. *Hum Mutat* **28:** 1171–1177. doi:10.1002/humu.20586

Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22:** 568–576. doi:10.1101/gr.129684.111

Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. 2013. The next-generation sequencing revolution and its impact on genomics. *Cell* **155:** 27–38. doi:10.1016/j.cell.2013.09.006

Koboldt DC, Mihalic Mosher T, Kelly BJ, Sites E, Bartholomew D, Hickey SE, McBride K, Wilson RK, White P. 2018. A de novo nonsense mutation in *ASXL3* shared by siblings with Bainbridge–Ropers syndrome. *Cold Spring Harb Mol Case Stud* **4:** a002410. doi:10.1101/mcs.a002410

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536:** 285–291. doi:10.1038/nature19057

Liu M, Jin M, Chen X, Wan B, Guo Y, Sheng M, Chen L, Zhao L, Huang D, Li Y. 2018. *ATP7B* mutation detection and pathogenicity analysis: one atypical case of Wilson's disease with adrenocortical insufficiency. *J Mol Neurosci* **64:** 20–28. doi:10.1007/s12031-017-0997-7

Loudianos G, Dessi V, Lovicu M, Angius A, Figus A, Lilliu F, De Virgiliis S, Nurchi AM, Deplano A, Moi P, et al. 1999. Molecular characterization of Wilson disease in the Sardinian population—evidence of a founder effect. *Hum Mutat* **14:** 294–303. doi:10.1002/(SICI)1098-1004(199910)14:4<294::AID-HUMU4>3.0.CO;2-9

Paz I, Kosti I, Ares M Jr, Cline M, Mandel-Gutfreund Y. 2014. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res* **42:** W361–W367. doi:10.1093/nar/gku406

Petrukhin K, Fischer SG, Pirastu M, Tanzi RE, Chernov I, Devoto M, Brzustowicz LM, Cayanis E, Vitale E, Russo JJ, et al. 1993. Mapping, cloning and genetic characterization of the region containing the Wilson disease gene. *Nat Genet* **5:** 338–343. doi:10.1038/ng1293-338

Petrukhin K, Lutsenko S, Chernov I, Ross BM, Kaplan JH, Gilliam TC. 1994. Characterization of the Wilson disease gene encoding a P-type copper transporting ATPase: genomic organization, alternative splicing, and structure/function predictions. *Hum Mol Genet* **3:** 1647–1656. doi:10.1093/hmg/3.9.1647

Roberts EA, Schilsky ML; American Association for Study of Liver Diseases (AASLD). 2008. Diagnosis and treatment of Wilson disease: an update. *Hepatology* **47:** 2089-2111. doi:10.1002/hep.22261

Shanmugavel KP, Kumar R, Li Y, Wittung-Stafshede P. 2019. Wilson disease missense mutations in *ATP7B* affect metal-binding domain structural dynamics. *Biometals* **32:** 875–885. doi:10.1007/s10534-019-00219-y

Tanzi RE, Petrukhin K, Chernov I, Pellequer JL, Wasco W, Ross B, Romano DM, Parano E, Pavone L, Brzustowicz LM, et al. 1993. The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. *Nat Genet* **5:** 344–350. doi:10.1038/ng1293-344

Umair M, Alfadhel M. 2019. Genetic disorders associated with metal metabolism. *Cells* **8:** E1598. doi:10.3390/cells8121598

Wang C, Zhou W, Huang Y, Yin H, Jin Y, Jia Z, Zhang A, Liu Z, Zheng B. 2018. Presumed missense and synonymous mutations in *ATP7B* gene cause exon skipping in Wilson disease. *Liver Int* **38:** 1504–1513. doi:10.1111/liv.13754

Yi L, Kaler SG. 2015. Direct interactions of adaptor protein complexes 1 and 2 with the copper transporter ATP7A mediate its anterograde and retrograde trafficking. *Hum Mol Genet* **24:** 2411–2425. doi:10.1093/hmg/ddv002